The Class Size Literature and What it Means for Illinois

David Figlio & Diane Whitmore Schanzenbach
Northwestern University

About us

Figlio is an economist and the Orrington Lunt Professor of Education and Social Policy at Northwestern University, and the Director of the Institute for Policy Research. His research focuses on education policies, especially regarding school accountability, choice, and finance. He is the immediate past president of the Association for Education Finance and Policy and the editor-in-chief of the *Journal of Human Resources*.

Schanzenbach is an economist and Professor of Education and Social Policy at Northwestern University, and Director of the Hamilton Project and Senior Fellow at the Brookings Institution. Her research focuses on education policies, especially the impact of school resources such as finances and class size on student outcomes. She is known as a class size expert, and has recently written the entry on "Class Size" for the *International Encyclopedia of Social and Behavioral Sciences* and the "Economics of Class Size" for the *International Encyclopedia of Education*, among other publications. She has also published six peer-reviewed papers using the Project STAR experimental class size data.

Overview

Class size is one of the most studied education policies, and an extremely rigorous body of research demonstrates the importance of class size in positively impacting student achievement. Illinois tends to have larger average class sizes than other states—the most recently comparable data ranked Illinois 36 out of the 46 states with data available, well above the national mean—and there is undoubtedly room to reduce class size as part of a school finance reform. That said, documents about the Evidence-Based Model overstate the case by inflating the relationship between class size and achievement, and by choosing an unrealistically low class size target that appears to be out of line with other states.

The research on class size

The research shows that students perform better in small classes. This is especially the case for students from disadvantaged backgrounds, who experience even larger gains than average students when they are placed in smaller classes. Small class sizes enable teachers to be more effective, and children who attend small classes in the early grades are permanently better off. We have reviewed many studies concerning class size and its impact on student outcomes, and summarize the findings below. Please note, however, that we only review studies that are capable of identifying the cause and effect relationship between class size and student outcomes. Such studies of the impact of class size on student outcomes must employ a research strategy that isolates the impact of smaller class size and does not conflate it with other factors. Studies that cannot separate correlation from causation are not included in our review.

Experimental evidence is strong for positive impacts of small classes in the early grades

The best evidence on the impact of reducing class sizes comes from Tennessee's Student Teacher Achievement Ratio (STAR) experiment.[1] A randomized experiment is the gold standard of social science research. In STAR, students and teachers in 79 Tennessee elementary schools were randomly assigned to small or regular-sized classes from 1985-89. The students were in the experiment during kindergarten through 3rd grades. Because the STAR experiment employed random assignment, any differences in outcomes can be attributed with great confidence to being assigned to a smaller class. In other words, students were not more or less likely to be assigned to small classes based on achievement levels, socio-economic background, or more difficult to measure characteristics such as parental involvement.

The results from STAR are unequivocal. Students' achievement on math and reading standardized tests improve by about 0.15 to 0.20 standard deviations from being assigned to a small class of 13-17 students instead of a regular-sized class of 22-25 students.[2] When the results are disaggregated by race, it appears that black students benefited more from being assigned to a small class than the overall population—approximately one-third of a standard deviation—suggesting that reducing class size might be an effective strategy to reduce the black-white achievement gap. Note that while there were no Hispanic students or English language learners in the STAR sample, in other settings researchers have found that education policy impacts are similar for black and Hispanic students.[3] Small-class benefits in STAR are also larger for students from low socio-economic status families, as measured by eligibility for the free- or reduced-priced lunch program.

Importantly, the positive impacts of small classes have been found not only on test scores during the duration of the experiment, but also on later life outcomes such as youth criminal behavior, teen pregnancy, high school graduation, college enrollment, college quality, college completion, savings behavior, marriage rates, residential location and homeownership.[4]

---

[1] Mosteller, Frederick (1995), "The Tennessee Study of Class Size in the Early School Grades," *The Future of Children*, 5(2): 113-127.

[2] Word, Elizabeth, J. Johnston, Helen Pate Bain, et al. (1990), *Student/Teacher Achievement Ratio (STAR): Tennessee's K-3 Class Size Study, Final summary report 1985-1990*, Nashville: Tennessee State Department of Education. Krueger, Alan B. (1999), "Experimental Estimates of Education Production Functions," *Quarterly Journal of Economics* 115(2): 497-532. Krueger, Alan B. & Diane Whitmore (2001), "The Effect of Attending a Small Class in the Early Grades on College Test-Taking and Middle School Test Results: Evidence from Project STAR," *Economic Journal* 111:1-28.

[3] E.g. Fryer, Roland G. (2010), "Financial Incentives and Student Achievement: Evidence from Randomized Trials," Harvard University Working Paper. Fryer, Roland G. and Steven D. Levitt (2006), "The Black-White Test Score Gap Through Third Grade," *American Law and Economics Review* 8(2): 249-281.

[4] Krueger, Alan B. (1999). Krueger, Alan B. & Diane Whitmore (2001). Krueger, Alan B. & Diane Whitmore (2002) "Would Smaller Classes Help Close the Black-White Achievement Gap?" In *Bridging the Achievement Gap*, eds. J. Chubb and T. Loveless, pp. 11-46, Washington, D.C.: Brookings Institution Press. Chetty, Raj, John N. Friedman, Nathaniel Hilger, Emmanuel Saez, Diane Whitmore Schanzenbach & Danny Yagan (2011), "How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project STAR," *Quarterly Journal of Economics* 126(4): 1593-1660. Dynarski, Susan, Joshua Hyman & Diane Whitmore Schanzenbach (2011), "Experimental Evidence on the Effect of Childhood Investments on Postsecondary Attainment and Degree Completion," NBER Working Paper

Note that the effect sizes reported in the Evidence-Based materials, ranging from 0.5 to 0.75, essentially double the effect size that we think is accurate.

An important concern for any experimental results is whether the results may be generalized to other settings. Along many measures, Tennessee in the mid-1980s looks similar to other places that might be interested in implementing a class-size reduction policy, so it would be reasonable to expect similar effects as those in the experiment. However, compared with the United States overall and with Illinois, Tennessee has lower levels of educational spending and lower teacher education levels. If adding resources has a greater impact when baseline levels are already low, it might suggest that schools with higher levels of spending could experience a smaller impact of class size reduction.[5]

Other high quality evidence lines up with STAR, and allows us to consider class size reductions of different magnitudes

Other high quality studies that isolate the causal impact of small class size in elementary school on student outcomes generally show results that are similar to those found in STAR. Wisconsin's Student Achievement Guarantee in Education (SAGE) program reduced pupil-teacher ratios in high-poverty elementary schools from between 21:1 and 25:1 to between 12:1 and 15:1. Small class attendance improved student achievement by approximately 0.2 standard deviations.[6]

Additionally, there are high quality studies using data from Israel, Denmark, Bolivia, and Sweden that find strong benefits from class size reduction in both math and reading scores, of a magnitude that is consistent with Project STAR's experimental results. The most famous quasi-experimental approach to studying class size reduction comes from Angrist and Lavy's use of a strict maximum class size rule in Israel and a regression discontinuity (RD) approach.[7] In Israel, Maimonides' Rule dictates that no more than 40 students shall be in one class. As a result, the maximum class size faced by a student drops changes dramatically when enrollment is close to multiples of 40. For example, if a grade has 80 students, then a school could offer only 2 classrooms, with 40 students in each. If a grade has 81 students, however, in order to abide by the maximum class size rule the school must offer (at least) 3 classrooms and maximum average class size falls to 27 students. Angrist and Lavy find strong improvements overall in both math and reading scores from smaller classes, of a magnitude that is consistent with Project STAR's experimental results. Consistent with the experimental results, they also find larger improvements among disadvantaged students.

#17533. Finn, Jeremy, Susan Gerber & Jayne Boyd-Zaharias (2005), "Small Classes in the Early Grades, Academic Achievement, and Graduating from High School," *Journal of Educational Psychology* 97(2): 214-223.

[5] Schanzenbach, 2007; Lazear, 2001.

[6] Molnar, Alex, P. Smith, J. Zahorik, A. Palmer, A. Halbach & K. Ehrle (1999), "Evaluating the SAGE Program: A Pilot Program in Targeted Pupil-Teacher Reduction in Wisconsin," *Educational Evaluation and Policy Analysis* 21(2): 165-77.

[7] Angrist, Joshua D. & Victor Lavy (1999), "Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement," *Quarterly Journal of Economics* 114(2): 533-575.

There have been several subsequent papers that have identified the impact of smaller class sizes using maximum class size rules in other settings. Urquiola uses a similar regression discontinuity approach in Bolivia and finds that a one standard-deviation reduction in class size (about 8 students in his data) improves test score performance by 0.2 to 0.3 standard deviations.[8] Browning and Heinesen find similar results on data from Denmark, even though average class size is much smaller in their study (20 pupils per classroom, compared to 31 students in Angrist and Lavy's Israeli data).[9] Most recently, Fredriksson et al. evaluate the long-term impact of class size using data from Sweden and measuring class sizes among students between ages 10 and 13 who were facing a maximum class size rule of 30 students.[10] In adulthood (measured between ages 27 and 42), students in smaller classes had statistically significantly higher levels of completed education, wages, and earnings. According to their findings, a one-student reduction in class size increases the probability of having a college degree by 0.8 percentage points—a larger magnitude than found in STAR when scaled similarly.

Another quasi-experimental approach is to use variation in enrollment that is driven by random fluctuations in cohort sizes across different years. Hoxby takes this approach using data from the U.S. state of Connecticut.[11] She finds no statistically significant positive effect of smaller class size, and the estimates have the statistical precision to rule out an effect as large as about one-fifth the size found in Project STAR. One drawback of the Connecticut study is that test scores are only measured in the fall, so the impact of the prior year's class size may be somewhat muted by time away from school in the summer. The discrepancy between these results and those of other well-identified experimental and quasi-experimental studies remains a puzzle.

Importantly, several of these studies investigate class size reductions among larger classes (e.g. from 31 to 23), so they help inform the discussion of threshold effects below. The studies from Wisconsin and Israel also find larger improvements among disadvantaged students.[12]

Do small classes matter in older grades?

Most of the high-quality evidence on class size reduction is based on studies of the early grades, and high-quality evidence on the impact of class size on outcomes in older grades is more

---

[8] Urquiola, M. (2006). Identifying class size effects in developing countries: Evidence from rural Bolivia. *Review of Economics and statistics*, *88*(1), 171-177.

[9] Browning, M., & Heinesen, E. (2007). Class size, teacher hours and educational attainment. *The Scandinavian Journal of Economics*, *109*(2), 415-438.

[10] Fredriksson, P., Öckert, B., & Oosterbeek, H. (2013). Long-term effects of class size. *The Quarterly Journal of Economics*, *128*(1), 249-285.

[11] Hoxby, C. M. (2000). The effects of class size on student achievement: New evidence from population variation. *The Quarterly Journal of Economics*, *115*(4), 1239-1285.

[12] Angrist, Joshua D. & Victor Lavy (1999), "Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement," *Quarterly Journal of Economics* 114(2): 533-575. Angrist, Joshua D. & Jorn-Steffen Pischke (2009*), Mostly Harmless Econometrics: An Empiricist's Companion.* Princeton, NJ: Princeton University Press. P. 267. Heinesen, E. (2009). "Estimating Class-Size Effects Using Within-School Variation in Subject-Specific Classes," *The Economic Journal* 120: 737-760.

limited. There is evidence that smaller class sizes in 8th grade positively impact test scores and measures of student engagement, and some evidence that these impacts are larger in urban schools.[13] Common sense suggests that class sizes can be larger for older students, though we have uncovered no studies that have credibly compared class size reduction across ages. (We would be happy to design a study to definitively test this conjecture in the Illinois context if there is interest; recall that the STAR experiment was entirely funded by the state of Tennessee.)

Is the impact linear? What is the "right" class size?

The best evidence to date comes from the STAR experiment, which estimated substantial positive impacts from class size reduction from on average 22 to on average 15. In addition, an influential 1979 meta-analysis conducted by Glass and Smith found strong impacts of class sizes below 20, but we hasten to point out that most of the studies reviewed in the Glass and Smith analysis do not meet the modern standards of empirical evidence that we require for our quality criteria.[14] Based on this, some researchers conclude that the evidence supports better outcomes only if classes are below some threshold number of 15 or 20. Sometimes the argument is extended to suggest that reducing class size is not effective unless classes are reduced to within this range.

In our judgment, the evidence supports the interpretation that the relationship between class size and achievement is linear. We base this conclusion on the pattern in the literature that points to similar per-student impacts of class size reduction, even across studies with a range of sizes for their "large" and "small" class sizes.[15] The broader pattern in the literature finds positive impacts of class size reductions using variation across a wider range, including class size reductions induced by maximum class size rules set at 30 (in Sweden) or 40 (in Israel). In fact, the per-pupil impact is reasonably stable across class size reductions of different sizes and different baseline class sizes. For example, when scaled by a 7-student class size reduction as in the Tennessee experiment, the Israeli results imply a 0.18 standard deviation increase in math scores which is nearly identical to the Tennessee results.[16] The weight of the evidence suggests that class size impacts might be more-or-less linear across the range of class sizes observed in

---

[13] Boozer, Michael and Cecilia Rouse (2001), "Intraschool Variation in Class Size: Patterns and Implications," *Journal of Urban Economics* 50(1): 163-189.
Dee, Thomas and Martin West (2011), "The Non-Cognitive Returns to Class Size," *Educational Evaluation and Policy Analysis* 33(1): 23-46.

[14] Glass, Gene V. & Mary Lee Smith (1979), "Meta-Analysis of Research on Class Size and Achievement," *Educational Evaluation and Policy Analysis* 1(1): 2-16.

[15] Angrist, Joshua D. & Victor Lavy (1999), "Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement," *Quarterly Journal of Economics* 114(2): 533-575. Krueger, Alan B. (1999), "Experimental Estimates of Education Production Functions," *Quarterly Journal of Economics* 115(2): 497-532. Molnar, Alex, P. Smith, J. Zahorik, A. Palmer, A. Halbach & K. Ehrle (1999), "Evaluating the SAGE Program: A Pilot Program in Targeted Pupil-Teacher Reduction in Wisconsin," *Educational Evaluation and Policy Analysis* 21(2): 165-77. Heinesen, E. (2009). "Estimating Class-Size Effects Using Within-School Variation in Subject-Specific Classes," *The Economic Journal* 120: 737-760.

[16] Angrist, J.D., & Pischke, J.S. (2009*). Mostly Harmless Econometrics: An Empiricist's Companion.* Princeton, NJ: Princeton University Press. See page 267.

the literature – that is, from about 15 to about 40 students per class. It would be inappropriate to extrapolate outside of this range. As a result, students in classes of size 24 or 25 are expected to be worse off in terms of academic achievement, student engagement, and the range of later life outcomes described above that have been tied to class size than students in classes of size 21 or 22, and a 3.5-student reduction in class size would be expected to have about half of the impact as the 7-student reduction seen in STAR and the related literature.

How does Illinois stack up?

Nationally comparable data come from the National Center for Education Statistics, and are somewhat dated (from the 2011-12 school year). In that year, Illinois ranked 35th for elementary schools and 36th for secondary schools out of 46 states with reported data.

	Elementary grades	Secondary grades
Illinois	22.9	27.7
US average	21.2	26.8
Illinois' ranking	36 out of 46	35 out of 46

Note that these are somewhat higher than the average class sizes reported in the state report card data, but have the advantage of being comparable across states. The Illinois report card data suggest that class sizes have not changed over the past 5 years, so it is probably safe to conclude that Illinois has relatively larger classes than other states.

Are there unintended consequences to dramatically reducing class size?

While the effects of reducing class sizes on student achievement are most likely linear, there is reason to believe that rapid wide-scale reductions in class sizes could undermine the positive educational benefits of class size reduction. A cautionary tale comes from the 1996 California class size reduction, where the sharp reduction in class sizes necessitated a large influx of new teachers, many of whom were inexperienced and certified on a temporary basis. This change in the teacher mix had the effect of dampening the benefits of the class size reduction.[17] Moreover, the schools where the benefits of class size were the lowest tended to be those serving large fractions of economically disadvantaged and minority students. In sum, it takes time for the supply of high-quality teachers in a market to catch up to demand, and in the short run, the schools serving particularly vulnerable students might find themselves in especially significant teacher shortages.

Is there a recommended class size?

The literature does not suggest that there is an optimal class size, nor that there is a special threshold under which class size reduction suddenly becomes effective. Instead, it suggests that class size reductions are cost-effective, in that their expected benefit on increased student test scores—and what that means for future earnings of students—outweigh their costs.

[17] Jepsen, Christopher and Steven Rivkin (2009), "Class Size Reduction and Student Achievement: The Potential Tradeoff Between Teacher Quality and Class Size," *Journal of Human Resources* 44(1): 223-250.

Aiming for a class size of 14 or 15 in the early grades is pretty far outside of the norm across states, as illustrated by the table of class size rules and average class sizes attached here. Even if these were the eventual goal for the state, we also would not recommend decreasing to these numbers quickly, because of the unintended consequence of teacher labor market responses to class size reductions.

It would certainly be reasonable to bring Illinois's class sizes down to meet or better the U.S. average, and we would expect this to improve student outcomes without breaking the bank.

Table 2	State Class Size Policies, Mandates and Voluntary Programs				
State	**Description**	**Funding Source**	**Waiver Available**	**Average Class Size (SY 11-12)**	**Year of Implementation /Adoption**
Mandate *Class Size Measured at Classroom Level*					
Alabama	K = 1:18 1-3 = 1:18 4-6 = 1:26 7-8 = 1:29	1995 Foundation Program Plan	Yes	19.4	1998
Delaware	K-3 = 22:1 or lower (classroom instructional aide counts as half of a teacher)	Not specified in statue	Yes	21.1	1998
Florida	Core Classes P-3 = 18 4-8 = 22 9-12 = 25	Legislature has appropriated >$25 B toward operational expenses and $2.5 B in facilities funding to implement	No	N/A	2003
Georgia	K = 18 1-3 = 21 4-8 = 28 9-12 = 32 (for core subjects)	State funding formula	Yes	21.2	2007
Hawaii	K-2 = 25:1 K-3 = 20:1 (Optimum class size) 4-12 = 26:1 (Optimum class size)	Not specified in rule	No	N/A	2004
Kentucky	K-3 = 24 4 = 28 5-6 = 29 7-12 = 31	Funding appropriated based on prior year's average daily attendance	Yes	23.7	1985

Louisiana	K-3 = 26 4-12 = 33 K-3 system-wide, student-teacher ratio 20:1	Not specified in statue	Yes	19.4	1986
Mississippi	K = 22:1 (27:1 w/full-time teaching aide) 1-4 = 27:1 5-12 = 33:1 (core classes)	funding reduced by % difference between actual and required pupil/teacher ratio	Yes	22.1	1990
Missouri	K-2 = 25 3-4 = 27 5-6 = 30 7-12 = 33	Not specified in rule	No	20.7	1990
Montana	K-2 = 20 3-4 = 28 5-8 = 30 (single grade classrooms)	Not specified in rule	No	20.5	1989
Nevada	K-2 = 16:1 3 = 18:1 alternative ratios for school districts in a county whose population is less than 100,000: 1-3 = 22:1 4-6 = 25:1	Not specified in rule	Yes	26.1	1989
New Hampshire	K-2 = 25:1, goal of 20:1 3-5 = 30:1, goal of 25:1 6-12 = 30:1	Not specified in statue	No	21.2	2005
New Jersey	K = 25:1 K = 21:1 (Abbott classrooms) K-3 = 21 ("at-risk" districts) 4-5 = 23 ("at-risk" districts) 6-12 = 24 ("at-risk" districts)	Not specified in rule	No (Abbott) Yes (non-Abbott)	19	2007

North Dakota	K-3 = 25 4-8 = 30 9 -12 = 30	Not specified in rule	K-3: No; 4-12: allowed max of 34 students/class for 3% of total classes	19.3	2000
Oklahoma	K-6 = 20	1017 Fund	Yes	21.1	1990
Tennessee	K-3 = 25 4-6 = 30 7-12= 35	New state funding formula for public schools; increased education funds via half-cent sales tax increase	Yes	17.8	1992
Texas	K-4 = 22	Not specified in rule	Yes	18.6	1995
Washington	K-3 = 17 4-12 = 25	Directs legislature to appropriate state funds deemed necessary to achieve class size	No	23.9	2015
West Virginia	K = 20 1-6 = 25 (may have up to 3 additional students per classroom)	Not specified in rule	Yes	19.2	1983
Mandate *Class Size Measured at Level Other than Classroom*					
Arkansas	K = 20:1 1-3 = 23:1 4-6 = 25:1, 7-12 = 30:1 Higher limits with T/A: K = 22:1, 1-3 = 25:1, 4-6 = 28:1 (measured at classroom level K & 7-12; district average 1-6)	Not specified in rule	Yes	20.4	1984

Maine	K = 20:1 1-8 = 25:1 9-12 = 30:1 (measured at school level)	Not specified in statue	Yes	17.8	1985
Massachusetts	K = 25 (measured as district average)	Not specified in statue	Yes	20.1	1968
Minnesota	K-3 = 17:1 (measured as district average in each grade)	State revenue allocated to districts according to funding formula	No	23.7	2001
New Mexico	K = 20:1 1-3 = 22:1 4-6 = 24:1 7-8 = 27:1 (English courses only) 9-12 = 30:1 (English courses only) (K & English measured at classroom level; 1-6 measured as average among grades)	Not specified in rule	Yes	20.5	1978; 1994 (grades 4,5, and 6)
North Carolina	K-3 = 24 (individual class measured at classroom level) K-3 = 21 (maximum average for all classes within the local education agency)	State board of education request funds in state budget request	Yes	19.8	1955
Ohio	K-4 = 1:25 (measured as district-wide average)	Not specified in rule	No	21.8	1968

South Carolina	K-3 = 30:1 1-3 = 21:1 (district average enrollment for reading/math) 4-5 = 30:1 (English language arts & mathematics) 4-5 = 35:1 (all other subjects) 7-8 = 35:1 9-12 = 35:1 (measured at classroom level and district average; average student-teacher ratio in any school cannot exceed 28:1 based on the average daily enrollment)	Districts receive funding under state and federal formulae; large portion through Education Finance Act	Yes	19.4	2007
Vermont	K-3 = 20:1 4-8 = 25:1 (measured as school average)	Not specified in rule	Yes	16.7	1997
Virginia	K = 24:1, no class being larger than 29 students 1-3 = 24:1, no class being larger than 30 students 4-6 = 25:1, no class being larger than 35 students 6-12 = 24:1 in English classes (measured at classroom level and as district average)	Not specified in rule	No; Yes if for experimental purpose in K-12 school	20.4	1988
Wyoming	Each district that maintains an average student-teacher ratio of 16:1 for grades K-3 is eligible for funds	School foundation program	Yes	17.4	2011
Large-scale Voluntary Programs					

California	K-3 = 20; legislation authorized formation of smaller classes and provided funding for those schools choosing to do so (measured at classroom level)	Class Size Reduction Program	N/A	25.4	1996
Iowa	K-3 = 17:1, allocates funds to achieve pupil-teacher ratio (measured as district average)	Iowa Early Intervention Block Grant Program	N/A	20.9	1999
Maine	K-3 = 15:1 (recommended),18:1(max.), localities can elect to adopt/receive funding (measured at classroom level)	Program funded via competitive grant program	No	17.8	1983
South Carolina	School districts which choose to reduce class size to 15:1 in grades 1-3 are eligible for funding (measured at classroom level)	Funding provided based on poverty index	Yes	19.4	1998
Virginia	K-3 = 24:1; financial incentives to reach maximum student-teacher ratio; if free lunch eligibility percentages is ≤16% school not eligible for funding (measured as district average)	Localities required to match state's share of payment based on composite index; funding calculated based on the % of students eligible for free lunch	Yes	20.4	1995
Wisconsin	K-1, 2-3 = 18:1 or 30:2 at the district/school's choice (measured at classroom level)	Schools receive state aid up to $2,250 for each eligible low-income K-3 child	Yes	20.8	1996